Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

CENSUS METHODS.

By Hon. E. Dana Durand, Director of the Census.

Accuracy is the fundamental requirement of all statistical work. This means, first, accuracy in collecting the original material; second, accuracy in compiling and tabulating it; and, third, accuracy in analyzing and interpreting it. The first and third stages in the process of statistical work are extremely difficult, and it is concerning these that I wish especially to speak. It is a matter of comparative simplicity to secure accuracy in compiling and tabulating statistics. The principal problem here is to secure economy of time and money. A few words only as to the methods of compilation and tabulation contemplated for the coming census will be sufficient.

Methods of Tabulation.

The population census of 1910, like that of 1900, will be tabulated by the use of punched cards. Every person in the United States is given a card on which the facts with regard to sex, race, age, birthplace, birthplace of parents, and the like are indicated by the punching of appropriate holes. The number of persons possessing each specific characteristic or combination of characteristics which it is desired to show in the final tables is then counted by means of electrical tabulating machinery, electrical contacts being made through the holes punched in the cards.

The punching machines to be used at the present census differ very radically from those used before, and will, it is believed, not only increase the rapidity of the work, but tend to reduce the number of errors on the part of the clerks doing the punching. With the old punching machine, if an error was made in a single item,—even though it might be the last item to be punched on the card,—the entire card had to be destroyed.

This consumed time, and also resulted in a tempation to the operator to let the error go uncorrected. With the machines now to be used, no hole is punched in the card until the keys for all the facts to be punched have been set, and, if the operator makes a mistake by depressing the wrong key, he can correct it before the card is punched. It was not found possible at the last census to check all of the cards back to the schedules, and it will probably not be feasible to do so at this census. Part of the cards punched by each operator are selected at random and compared back, and, if any appreciable percentage of error is discovered in those thus compared, the other cards punched by the same operator are likewise compared.

The tabulating machines to be used at the present census will likewise result, it is believed, in a material increase of rapidity and reduction of errors. This is principally brought about by the fact that the results of the count for each unit of area are automatically printed, whereas formerly they were registered on dials from which readings had to be taken and recorded by hand. The reading of these dials took a large amount of time, during which the machine was idle, and inaccurate readings were not uncommon.

THE MARGIN OF ERROR.

Turning now to the more important subject of the means of securing accuracy in the original census returns collected in the field, it is self-evident that the entire value of the census depends on the securing of approximate accuracy in the original returns. Absolute accuracy is out of the question, and, in fact, a small margin of error does not seriously affect the value of the statistics; but any considerable margin of error practically destroys their value. No degree of accuracy in tabulating and no degree of skill and judgment in analyzing statistics can give value to data which were incorrect in the first place.

It has too often been a vice of statisticians to present to the public tables purporting to show all sorts of important facts without due consideration of their truthfulness. The general public who use the statistics—in fact, even the trained statisticians who use them—have in most cases no means of discovering the inaccuracy of such statistics and go on using them in misguided confidence. The compiler of statistics has often been, unconsciously perhaps, careless as to their accuracy because of the lack of any possibility of detection. The man who has charge of the collection of statistical data ought himself to be the severe and uncompromising critic of those data. He alone has approximately adequate means of judging the degree of accuracy which has been secured; and it is his duty, having done all that is possible to eliminate error, to inform the public fully and frankly of the extent to which error presumably still persists. So far from taking advantage of the fact that others cannot discover the errors which are hidden away in imposing-looking totals, he should from that very fact recognize the more clearly his own duty to explain just how much or how little the statistics may be trusted.

This does not mean, of course, that the statistician who knows that there are certain errors in his figures should straightway declare them of no value. Statistics on certain subjects may be of very little value unless almost absolute accuracy is secured, but there are many subjects as to which close approximations to the truth are almost as valuable as the exact truth itself. It is the mark of the competent statistician to be able to decide approximately what the margin of error actually is, and also to what extent the error vitiates the results. As Josh Billings said, "It is better not to know so much than to know too many things what ain't so."

At the census of 1900 and during the so-called intercensal period since that time the Census Bureau has done a great deal in the way of criticising its own statistics in the manner I have suggested; but I doubt whether it has yet been sufficiently emphatic in cautioning the public with respect to the existence of unavoidable errors. As one means of expressing the approximate character of the data, I think it desirable to express all very large totals of manufacturing and agricultural statistics in millions, and, in the case of the smaller figures which appear in the statistics for individual localities and industries, to express the results in thousands.

To criticise statistics, however, is easy: the difficult task is to improve them. The Census Bureau will have done part of its duty to the public if it gives warning regarding the margin of error in the statistics it publishes, but it is far more important that it should reduce that margin of error.

METHOD OF SELECTING FIELD FORCE.

There are two means to this end. The first and more difficult is to secure better men to collect the statistics: the second is to simplify and clarify the inquiries.

As you are aware, the statistics of population and of agriculture are collected by a different force from that employed in gathering the statistics of manufactures. The population and agricultural data are secured by enumerators, of whom there will be about sixty-five thousand at the present census, they in turn being appointed by the supervisors, of whom there are about three hundred and thirty. The difficulty of securing competent and faithful enumerators is very great. The length of service is short, fifteen days in the cities and thirty days in the country districts. The period is thus too short to justify a man who has a good job in quitting it, while, on the other hand, it is too long in most cases to enable such men to get leave from their regular work to take the census. Moreover, the pay is small, averaging perhaps $3 per day in the country districts and a trifle more in the cities,—practically the pay of ordinary mechanics. Most of those who seek, from financial motives, to be enumerators are men who are able to command only moderate pay in their occupations, while many of them are men who cannot command regular employment and who are looking for odd jobs.

Consideration has been given by the census authorities from time to time to the plan pursued in Germany and some other European countries, by which the census is taken chiefly or wholly by men serving without pay, who either volunteer their services from patriotic motives or who are required to act. I very much doubt whether conditions in this country are ripe for such a scheme. At any rate, nothing of this sort can be done at the present census. It may, however, be hoped that to some slight extent, and possibly to a considerable extent, men can be induced to accept the position of enumerator from interest in the work rather than for the compensation involved. I hope that a considerable number of the colleges and universities of the country may see fit to give leave of absence to their students for the short time required to do this work of enumeration. It is, however, usually desirable that enumerators should actually live in the districts where they work, and there are multitudes of districts where no college students reside or where such students are in institutions hundreds or thousands of miles from their homes. Another class who can render good service as enumerators are school-teachers, but, with the enumeration taking place in April and May instead of June as formerly, few school-teachers can be spared from their duties to take the census.

The primary responsibility for securing efficient enumerators must rest with the supervisors of the census. It has been suggested from time to time that more efficient enumerators might be secured by competitive civil service examination. The difficulty with such an examination is the expense and delay involved. It is probable that for the sixty-five thousand places there would be several hundred thousand candidates, and the grading of their papers would require a large force for a long time. At some future census this plan might be worth a trial, but it can scarcely be attempted at the present census within the limits of time and appropriations set by law.

The Census Bureau does, of course, undertake to protect itself against such obviously incompetent enumerators as the supervisors, through political influence or through oversight, may happen to choose. At the last census the enumerators recommended by the supervisors were all subjected to a test examination, and the same policy will be pursued this time.

About one-sixth of the candidates selected by the supervisors were rejected as the result of this test in 1900.

I hope and believe that the supervisors at the present census are, on the whole, a higher type of men than those at any preceding census. The compensation offered to supervisors is somewhat more than ever before, but it is not really an adequate remuneration for men of the character needed. I believe that a very considerable proportion of the supervisors who have been appointed are men who in their regular occupations or professions are able to earn much more than the supervisors' pay, and who have accepted the positions because of the honor and responsibility involved or from patriotic motives. There has, however, been no new departure with respect to the general method of selecting supervisors, save only that in most of the large cities, selections have been made independently of political recommendations.

It may not be inappropriate here to explain the reason why so-called political recommendations must be sought in the appointment of supervisors in most of the districts. The supervisor, in order to do his work properly, must be a resident of the district over which he has charge. The position is not one which can be filled by civil service examination. A supervisor should be a man of executive ability, such as no examination can test, and he should be a man well known in his community. Indeed, men of the type desired would not in most instances be willing to submit themselves to a civil service examination. Neither the Director of the Census, his superior officer, the Secretary of Commerce and Labor, nor the President, can possibly have personal acquaintance throughout the country with men suitable to fill these positions, nor even with men, outside of political positions, whose advice they can seek. The representative of the district in Congress or the party leaders in the district have the necessary knowledge of local conditions and local men, and it is not only natural, but practically necessary, as a matter of good administration, that they should be asked to make recommendations for positions of this character.

Fortunately, members of Congress and party leaders are coming more and more to appreciate the importance of accurate census statistics. The spoils idea of appointments is growing constantly less dominant. The Director of the Census has done everything possible to impress upon those whose opinions were asked regarding the supervisorships the importance of the work and the necessity of having men of marked executive ability and of absolute integrity. The information furnished by members of Congress and party leaders regarding the candidates whom they have recommended has been supplemented by full personal statements required from each candidate, and in most cases by numerous letters of indorsement from business and professional men.

In the large cities the administration is evidently less dependent than in the country districts upon members of Congress and party leaders for information regarding available material for such positions as that of supervisor of the census, and the President and the Secretary of Commerce and Labor have joined with the Director of the Census in establishing the policy of making selections, wherever the information on which to act was available, without depending on the recommendations of political leaders.

The work of collecting statistics of manufactures and mining is in some respects even more difficult than that of collecting population and agricultural statistics. The number of employees required, however, is much more limited, and consequently it is possible to employ more elaborate methods of selecting them. The manufactures and mining statistics are practically all collected by men known as special agents. These are appointed by the Director, and the law does not require any examination, but it is my intention, nevertheless, to make the appointments as the result of competitive examination, held on November 3, from which, however, men who have had previous successful experience in field work for the Census Bureau will be excused, as well as men who have successfully passed the somewhat similar examination for the position of special agent in the Bureau of Corporations. It is evident that an ordinary civil service examination, with its tests merely of what a man knows, is not adequate for the selection of men to do the responsible and largely independent work of visiting manufacturing and mining establishments to secure schedules from them. Consequently, following a precedent set by the Bureau of Corporations and to some extent employed also in other more advanced scientific civil service examinations, the examination for special agents consisted in part of the presentation of evidence regarding the applicant's previous education and experience, while the remainder of the examination was a very practical test of his ability to fill out a manufactures schedule. For those who are to do the simpler field work, this test was little more than reproducing on a schedule a set of facts which have been stated in narrative form, but, for the agents who are to visit the more important establishments or who are to have charge of other agents, the test consisted of the filling of a schedule from somewhat complicated details in a hypothetical balance sheet, income account, and pay-roll. Moreover, not being bound by any legal requirement to take the man with the highest rating first, the appointing officers of the Bureau will be able to give proper attention to such matters as personal address and tactfulness, which cannot be tested by any examination. We hope to secure as special agents largely young college and university men with some training in economics and statistics, or men who have been book-keepers of the more advanced type in manufacturing establishments or who have had other direct business experience.

PREPARATION OF SCHEDULES.

The form of the schedules on which census statistics are collected has an extremely important bearing upon the accuracy of the results obtained. In the drafting of a schedule, the first consideration is naturally what information it is desired to secure, but it is equally important to consider how far it is possible to obtain information accurately. There are always abundant suggestions from people interested in different economic and social problems as to points which ought to be included in the schedules, and there is comparatively little danger that a statistician in public office will neglect to make inquiries merely because he does not know that the information would be of utility to the public. There are not so many people, however, who are well informed as to the difficulties of securing statistical facts; and one of the most important tasks of the statistician is to determine what facts can be secured with accuracy by the means which he has at his disposal. Not only are some inquiries, which, if answered correctly, would furnish very valuable information, in themselves quite incapable of correct answer under existing conditions, but the multiplication of inquiries which separately are quite capable of accurate answer tends to lessen the accuracy of the replies to all of them. A schedule may, in other words, easily be overloaded. The enumerator or special agent collecting the statistics becomes discouraged if too much is asked of him, and the person to whom he addresses his inquiries becomes still more discouraged, not to say annoyed and confused.

Every reduction in the number of inquiries contained in a census schedule tends to greater accuracy in the returns, except of course in those cases where such reduction in the number of inquiries injures the clearness of the schedule. Some details in schedules, themselves comparatively unimportant, are often necessary in order to make clear what is intended to be included in a total.

This demand for reduction of the number of inquiries and simplification of them, in the interest of accuracy in the returns themselves, is in some cases opposed by the likewise important demand for details which will permit an accurate analysis and interpretation of the statistics. In other words, the lower form of accuracy sometimes stands opposed to the higher form, which seeks the true spirit of the phenomena investigated. For example, take manufactures statistics. In many industries the entire technical and economic conditions are rapidly changing, and broad general interrogatories, even though answered with entire accuracy, give results which are not comparable with those of preceding censuses because of such changes in conditions. One instance of this character happens to have come to my personal attention recently. The schedules for the petroleum industry call for the value of the different kinds of oils produced, including the containers in which they are shipped from the refineries. Twenty or thirty years ago nearly all oil was shipped from refineries in barrels or other small containers, the value of which represented a large fraction of the total value of the product. At the present time the larger part of the products of refining is handled in bulk, without containers. A comparison of the average value computed by dividing the number of gallons into the total value of the product, including containers, from one census period to another, is, therefore, highly misleading. The higher form of statistical accuracy would be secured only by the insertion of somewhat burdensome questions, designed to separate the value of the product contained from that of the containers.

It is evident, therefore, that it is an exceedingly important and difficult task to frame a schedule which shall combine simplicity with that completeness which will permit correct interpretation.

The manufactures, mining, and agricultural schedules are necessarily the most complicated which the Census Office employs. The office has spent a very large amount of effort during the present summer in studying these schedules, and has called to its assistance a number of economists and statisticians from the universities, as well as men connected with the practical work of the Department of Agriculture and the Geological Survey and men actively engaged in business. I hope that we are succeeding in materially reducing the burden upon the enumerators and the special agents, while at the same time cutting out few inquiries of material value and adding some which will make the results more intelligible.

To illustrate what has been undertaken, I will call attention to two very material changes which have been made in the schedule of manufactures. At the censuses of 1900 and 1905 that schedule called for the average number of men, the average number of women, and the average number of children under sixteen employed during each month in each establishment. To answer this interrogatory with strict accuracy would have required the special agent or the manufacturer to go in detail through every pay-roll of the year, usually either fifty-two or twenty-six in number, counting the number of men, women, and children separately on each,—the segregation of those under sixteen involving peculiar difficulty,—and averaging the account for each month. To do this and at the same time to answer the inquiry regarding classified wages to which I will shortly refer would, in my opinion, have required ten times as much time as to answer all of the other inquiries on the schedule put together. As a matter of fact, practically none of the returns were, I am convinced, based on analysis of the pay-rolls. They were in nearly all cases mere estimates made more or less off-hand by the manufacturer or his book-keepers, and, as many months had elapsed before the estimates were made, there is every reason to believe that they were often very wide of the mark.

The schedule as it has been revised for the present census calls for the number of men, women, and children employed during one specific week only, and for the total number employed, without distinction of age or sex, on the fifteenth day of each month. The relative extent to which women and children are employed will be ascertained approximately from the figures for the selected week. If the total number of employees for one day of each month is returned with approximate accuracy, then the average of these twelve dates will unquestionably show very closely the true average number of employees during the year. We hope to be able to induce special agents and manufacturers in a large proportion of cases to answer this revised interrogatory from actual payrolls, and, if so, the greater accuracy will far outweigh any apparent sacrifice in the amount of information obtained.

At the last census of manufactures a question was included in the schedule which called for the number of employees classified according to their earnings for one week of the year, separating men, women, and children under sixteen. This is a question which, if correctly answered, would furnish information of very great value. We have become convinced, however, that for many establishments correct answers cannot be obtained, and that for other establishments the amount of effort required to obtain correct answers is so great as scarcely to justify the expense, while also tending to injure the accuracy of all the data called for on the schedule. Strictly speaking, this inquiry falls within the field of the Bureau of Labor rather than that of the Census Bureau. It may be, however, that at another census, in order to take advantage of the fact that census agents are to visit every manufacturing establishment, it would be desirable to ask the authority of Congress to provide a special schedule covering classified wages, and an extra appropriation to cover the expense of getting it properly filled from actual pay-rolls. It would not be essential that returns of classified wages be secured from every manufacturing establishment. If such returns could be secured accurately for even 50 per cent. of the employees, they would be of very great value, and the agents could be specifically instructed to accept no estimates whatever.

What has just been said regarding these two changes in the manufactures schedule will illustrate what I said before with regard to the increase in accuracy which can be secured by simplification of the schedules. The schedule of manufactures hitherto has been overloaded. Since special agents and the manufacturers who filled them were practically forced in many cases to make estimates in replying to the interrogatories regarding the average number of employees and classified wages, they were tempted to regard estimates as sufficient for the replies to all the interrogatories of the schedule. The intellectual integrity of the special agents was undermined by requiring of them work which they knew could not be done accurately within the time which they were allowed.

ANALYSIS AND INTERPRETATION OF STATISTICS.

It is a commonplace of statistical science that the mere presentation of figures has comparatively little value without analysis and interpretation.

The statistician owes to the public the duty of properly analyzing and interpreting his material, not simply in order that the public may get the full value of the results, but, what is more important, in order that false and misleading interpretations may be forestalled and prevented. Whether the untruthfulness has been due to intentional misrepresentation or to ignorance, census statistics have probably been made the basis of more fundamentally untruthful speeches and articles than any other set of facts. The reports of the Census Bureau in the past have contained many warnings against such misleading use of their contents, but such warnings are all too often disregarded. Liars and fools—to use perhaps exaggerated expressions—will continue to abuse census statistics, whatever the Government may do to prevent it. We shall attempt at the present census, however, to make the warnings against such use of the figures even more numerous and more emphatic than in the past, in order that the man of ordinary intelligence may in general escape the pitfalls. Too often the warning has been posted in an inconspicuous place, hidden away in some long page of text, when it should properly have been presented in the most conspicuous place in the text, or often in the headings of the tables themselves, since most people, even those of considerable economic training, usually pay little attention to the text and assume that the tables mean what on their face they purport to mean. These warnings should relate not merely to unwarranted conclusions from the figures, but also, where necessary, should caution the reader against assuming a degree of accuracy in the statistics which does not exist. One of the most important duties of the census statistician is, as I stated at the beginning, to discuss frankly and clearly the margin of error in the returns.

In addition to increasing the frequency and the emphasis of the warnings against improper use of the census statistics, it will be desirable, in my opinion, to extend somewhat further than hitherto the analysis and interpretation of the statistics, at least in certain directions. A desirable feature of this increased analysis will be the publication of additional monographs dealing with particular economic and social problems. Many people have little interest in census statistics in general, but are profoundly interested in the data regarding some particular subject, and such data should be available for their use in separate form so far as possible. You will recall, for example, the exceedingly able monograph prepared by Dr. Willcox on the basis of the census of 1900 and dealing with the negro. We contemplate another monograph on this subject at the present census, and also similar monographs dealing with the foreign-born, with the family, with occupations, and perhaps with other similar subjects connected with the population statistics. Monographs dealing with particular branches of agriculture are also contemplated. The practice of presenting the statistics of the leading manufacturing and mining industries in monographic form has already been pursued quite generally at previous censuses.

Proper analysis and interpretation of the census statistics will call for the employment of a considerable number of statistical experts in addition to those on the permanent roll of the census. A considerable proportion of the tabular and text analysis and interpretation of the statistics at the previous censuses has been committed to men of comparatively little economic and statistical or business training and experience, although beyond question much of the work has been done by highly competent men. We hope to be able to enlist the service, for periods ranging from a few months to a year or two, of a considerable number of expert special agents to aid in the work of analyzing and interpreting the returns. Hitherto the limit of salary for temporary expert service of this sort has been altogether inadequate, $6 per day. In July of this year Congress passed a bill authorizing the payment of not to exceed $8 per day to twenty expert special agents. The Director had asked that the limit of pay be fixed at $10, and it is possible that Congress will this winter be again urged to establish that rate. Even $10 per day is but moderate compensation to offer to men who possess the necessary qualifications for this work, most of whom are already earning at least that much in permanent positions.

It is not the intention to employ exclusively men who have had university training in economics and statistics in this advanced work of analyzing and interpreting the census material. Particularly in the case of the manufacturing, mining, and agricultural statistics, it will be desirable to obtain the service of men of practical business experience. This policy has been pursued, in fact, to a considerable extent at previous censuses. I believe, however, that in some cases the economy as well as the scientific value of this work has been impaired by having too much of it done by men who were not actually present in Washington, but who gave only such time as they could spare from their regular professions or occupations to the work at their own homes, coming only occasionally to Washington for consultation. This practice has tended to undue lack of uniformity in methods, and has placed too great responsibility on single individuals. While specialization has its great advantages, it should be combined with consultation with others working on kindred topics.

The general policy of insisting that, for the most part at least, the expert special agents who analyze the statistics and prepare the text of the reports shall be present in Washington at least the greater part of the time, will in some cases exclude practical business men such as have formerly been employed for that purpose, because of their inability to give up, even temporarily, their regular duties. We shall, however, employ such men as consulting experts, having them come occasionally to Washington or sending the expert special agents who are assigned to particular tasks to consult with them at their homes. We believe that the undivided attention of men of high capacity and of general statistical and economic training, even though they may previously have had little direct experience in the particular subject intrusted to their charge, will enable them in a comparatively short space of time to become sufficiently expert to handle the statistics even more satisfactorily than can be done by men of more practical experience who are not personally present and can give only a part of their time.